Fundamentals of Multimedia Computing

Chapter 22: A Note on Privacy and Security Issues

Authors:
Gerald Friedland
and
Ramesh Jain

Draft for Comments

A Note on Privacy and Security Issues

In this chapter, we will discuss some issues that are not so much of technical nature but are a side effect of the power of multimedia. Since multimedia data is directly encoded for human perception and does not necessarily undergo a creator's mind filter (like text), there are some problems that do not usually arise with the creation, distribution, and consumption of other types of data.

Issues of the Effect of Multimedia Data

Probably the most commonly discussed issue of multimedia data is that of the effect of multimedia content on people. Technically, humans could be interpreted as sensors perceiving the data encoded and then interpreting it. This can lead to emotional reactions ranging from entertainment and excitement to sadness and even trauma. The reason for this is that multimedia data allows us to perceive people, objects, events, actions, and places that we would not be able to perceive without multimedia recording.

Naturally, this is foremost a problem in children. Therefore, magazines, movies and computer games are usually rated for a certain age group. While this is true for books as well, it's not as strongly enforced in most places -- another indication for the power of direct encoding for perception. News broadcast usually do not show scenes that are too cruel in order to not expose the audience to material that they may have never experienced and that potentially would lead to trauma and other negative psychologic effects. Some material, such as pornographic videos, is directly targeted to evoke certain emotions and anatomic reactions.

The creation and distribution of certain multimedia material is forbidden in almost all countries. While different countries have different rules and reasons to prohibit handling different types of multimedia data, the cause usually is the same: Society fears the power of multimedia data to directly manipulate people and is still researching long-term effects.

Multimedia data can also cause physical harm. The easiest example is audio that is too loud or too highly pitched. Ear damage is caused by people listening to ear phones too loudly. While this seems almost obvious, it is very important that any multimedia system has to have the ability to be shut down quickly, e.g. with a pause button. When presenting sound, a user friendly method of regulating the volume should definitely provided. People should not be exposed to loud pop sounds and/or be surprised by a sudden increase in volume. Pitch also plays a role. Certain tones can be perceived as uncomfortable when the audience is exposed to them for too long. Also, there is a myth that low-frequency noise (around 10 Hz) can cause nausia in people (see references). While we could not find literature to validate the myth, this at least shows that people intuitively are not sure about the exposure to certain sounds.

Visual data is also able to do harm. Spending too much time concentrating on low-contrast content, e.g. reading yellow font on white background, will usually be reported as uncomfortable and is often attributed to be a cause for headaches. Stroboscopic light, for example induced by flashing monitors, is reported to cause epileptic seizures in some persons. Many computer games warn about this issue. For the same reason, the Web Content Accessibility Guidelines (WCAG)

Version 2.0, produced in 2008, specifies that content should not flash more than 3 times in any 1 second period.

3D video and augmented reality applications are sometimes the cause for sea sickness symptoms. The reason for this is that our visual perception does not exactly match our proprioception and other sensors and the reaction varies from person to person: From no reaction to dizziness and nausia. The effects of 3D displays on young children are currently discussed: Since visual perception may not have fully developed yet, consuming artificial 3D data may cause perceptual issues due to the brain adapting to wrong realities.

Consuming multimedia data can also cause unwanted and potentially life-threatening distractions, especially when operating machinery or driving a car. While it is generally acknowledged that listening to music on moderate levels while driving is not harmful, listing to music using earphones is prohibited in many countries as the music would mask acoustic events from the outside and therefore reduce the ability of the driver to react to them. Watching a movie while driving would make it nearly impossible to drive a car, as the split attention effect (see chapter XXX) would slow down reactions to traffic to a very dangerous level even at low speeds. Operating a cell phone or a navigation system while driving has a similar negative effect.

It is possible to manipulate a human being in a very subtle way and potentially cause deep emotions (wether planned or not). Therefore it is very important for multimedia researchers to design systems in a way that it obeys the laws of different countries when handling multimedia data globally and make sure to prevent any harm from the people exposing the data to.

Privacy Issues

Multimedia data, especially acoustic and visual data, but in general data that is directly encoded for perception has the property of always conveying more information than is intended. The reason for it is the information richness: In contrast to written text, multimedia captures a snapshot of reality unfiltered. Even in staged scenarios, such as movies, people find bloopers that were not intended to be shown and were not part of any script. Even more so, spontaneous recordings such as vacation photographs of home videos always show lots of information that was not in the focus of the attention of the creator but is still there. This must be taken into account when publishing photographs, audio, and video recordings on the Internet or elsewhere.

Especially the wide spread use of social Internet sites to dissiminate multimedia materials has been questioned as problematic. First, many recordings may contain people and.or objects that require permission to be published. In many countries publishing a face photograph of a third person requires permission of that person. Copyrighted objects, such as a painting in the background, require permission of the author or current copyright holder.

Also, both human intelligence as well as multimedia retrieval may be used on multimedia data to extract some of the unintended information. For example, face recognition or speaker identification (see chapters XXX) may be used to find the identity of a person, even if the website has no other indication for a person's identity. This could potentially de-anonymize website postings and accounts. For example, a person having videos and photos on an otherwise anonymous might be identified by the picture or voice by people who know them. Furthermore,

multimedia retrieval, even though not perfectly accurate, makes it unsafe to say that data would be anonymous on the site: Only one match of a face or voice between the anonymized site and a public site with identity is enough to expose the user.

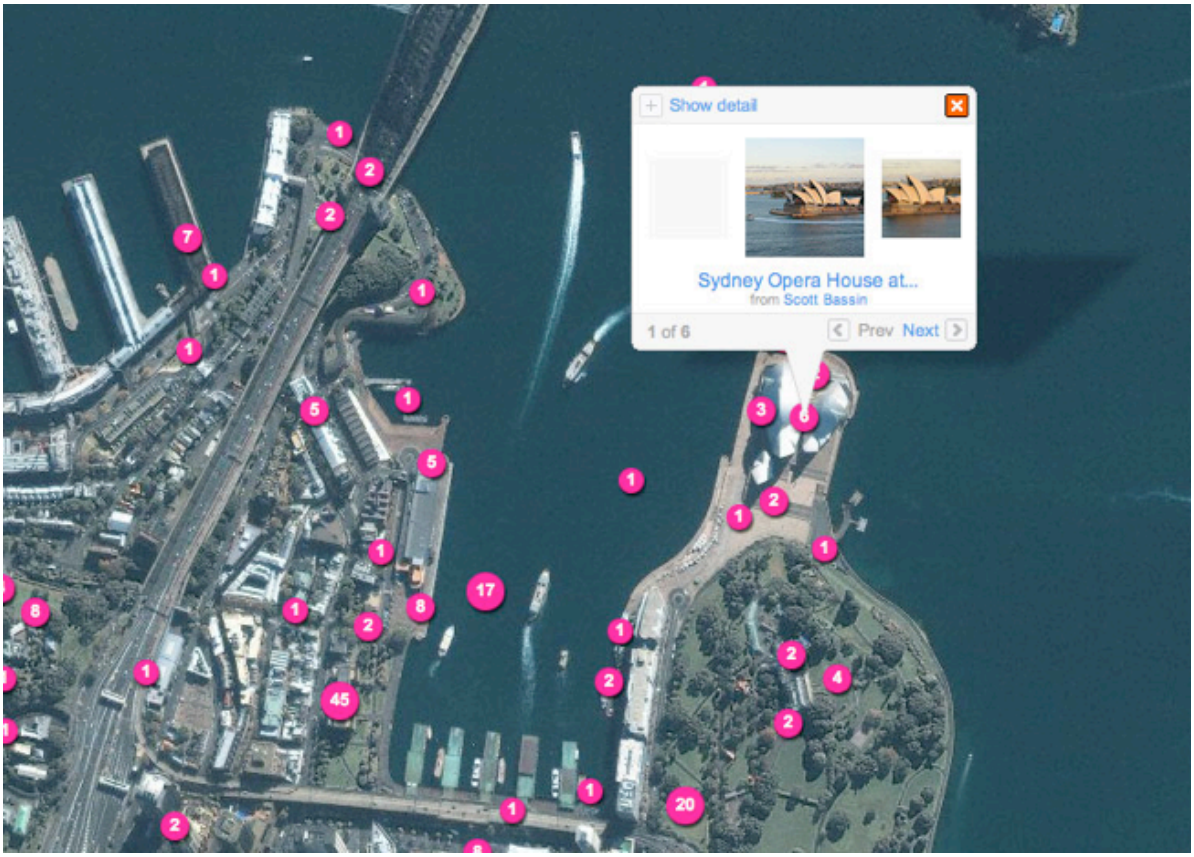

Figure 1. An example of geo-tagging allowing for easier search and retrieval of images. The photo shows a partial screenshot of flickr.com.

Even when such subtle methods like the ones described above are not in question, metadata and annotation might become an issue too. For example, tagging somebody publicly in a picture makes searching the person much easier as textual search has higher accuracy than multimedia search. However, this implies that a photo of a person has become public without the person potentially not knowing about it. This can be especially problematic because the person in question might have opted to not participate in the social networking site and is now forced to do so. Often, cameras and multimedia editing tools embed metadata in videos and images. The most common being the so-called EXIF header in JPEG images. EXIF can contain detailed information about the camera, including serial numbers, which potentially makes the creator of the photo trackable. Most importantly, EXIF data can contain geo-tags, i.e. longitude and latitude coordinates of the place where the photo has been taken. Figure 1 shows an example. This is enabled through different localization systems, such as GPS, cell-tower information, as well as wireless network SSID-maps. Geo-tags in combination with time allow a reasonable easy tracking of a person. Also, implicitly exposing places, such as home addresses, can be potentially

dangerous as it allows a variety of crime scenarios. Often, taking a photograph of an object is a matter of seconds and the creator is often not thinking about metadata, such as GPS tags. As a result, valuable items have been posted on the web including a complete address, whilst the post was anonymized otherwise (see references).

Since a potential attacker might gather information from different websites and infer from the collective, it becomes even harder to track what information is out there about an individual or entity. So conserving privacy becomes a difficult task, once a considerable amount of information has been published.

Time is another variable in the equation. What people might find adequate to post and publish now, might later be an isse for their reputation. The need for privacy shifts with age, social status, occupation, and other factors. On the other hand, the Internet potentially saves material forever and in many copies. Once data has been published, it might be downloaded by various search engines for caching and individuals might be wanting to re-post it. Seemingly inocous information might later become embarrasing whether posted intentionally or unintentionally.

As multimedia researchers, we have special responsibility in this regard as we are the original enables of most of the technology. Many people are unaware that multimedia retrieval technologies exist and even many experts often fail to assess the capabilities of today's multimedia retrieval technologies and the potential inference possibilities. Even though current retrieval methods are not perfectly accurate, with billions of pictures and millions of videos available, even a seemingly small fraction of true positive results can already translate into several hundred relevant matches. Those matches might enable de-anonymization, finding potential victims for crimes, or reveal other secrets.

Countermeasures

The first thing to do is checking whether a particular post is really necessary. What is the target group of the post and is it possible to make the post only available to that group.

Then, the most important privacy-conserving counter measure is to make sure only information is published that is supposed to be published. When publishing photos, for example, these should not contain humans that were not involved and if so, faces and other identifying characteristics should be blurred. Likewise, multimedia data should be checked for hidden metadata. If metadata is to be included, the level of detail should be controllable. For example, serial numbers of the camera might not be actually needed and geo-tags might be included with a reduced accuracy that is enough to organize the data but not enough to pinpoint individual addresses. Inserting noise into the signal which might or might not be audible might help to make it more difficult to apply acoustic matching. Applying compression often helps to reduce camera artifacts.

The above are only a few examples of both privacy issues and countermeasures. Spotting the issues and inventing new methods to prevent them is an active area of research and the reader is therefore suggested to check out the literature.

Security Issues

Apart from what has been said about the possible effects of the consumption of multimedia data in humans, multimedia data has also often attributed to computer security issues. Malware has been embedded in various forms into videos and presentations. This is mostly enabled through the fact that some multimedia data formats are turing complete, i.e. their representation is powerful enough to allow the execution of arbitrary programs. Prominent examples of these formats are Postscript and Flash, the first being a printer language, the second being an interactive video and animation format. PDF, that replaces postscript, is therefore not a Turing complete programming language anymore.

Security breaches have also been reported, exploiting buffer overruns in multimedia formats. Often, decoders and viewers of multimedia data are tailored to and tested only with regular images and videos. However, the nature of compression formats makes it possible to artificially create a video or image file that expands into a super large file upon decompression. The memory used to hold that overlarge data chunk potentially overwrites code, which then leads to crashes, or when done carefully could be used to insert malicious code into the decoder/viewer. Several viruses have been reported doing that. These viruses are especially powerful because they spread when a person views an image or video embedded in a webpage or email, thereby not even consciously executing a program.

Reportedly, there have been programs using microphones and cameras built into laptops without notifying the user. This spy-behavior inadvertendly leads to trust issues as the user should be always informed about the fact that he or she is being recorded. More so, it should be clear what is being done with the recording and how to stop the recording at any time. A video feedback helps improve a user's trust in what's being recorded and or transmitted. Needless to say, in many countries recording a person without their knowledge is against the law.

Again, this chapter only provides an introduction to the most pressing topics in the field. Privacy and security issues are hard to convey comprehensively a book as they are usually a natural side effect of a new technology. However, whenever creating a new technology we ought to ask ourselves about the impact on privacy and any potential security issues that may raise. Everybody working in the field should develop a common sense for these issues and not hesitate to question methods that might have unwanted side effects.

Literature

- Durham, M. & Kellner, D. (2001), *Media and Cultural Studies*. UK: Blackwell Publishing
- *Fowles, Jib (1999), The Case for Television Violence*, Thousand Oaks: Sage
- *Gauntlett, David (2005), Moving Experiences - Second Edition: Media Effects and Beyond*, London: John Libbey
- Anderson, C. A. & Bushman, B. J. (2001) Media Violence and the American Public: Scientific Facts Versus Media Misinformation. *American Psychologist*
- Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann and W. Prinz (Eds.) Relations between perception and action. Berlin: Springer.

- Mussen, P., & Rutherford, E. (1961). "Effects of aggressive cartoons on children's aggressive play" *Journal of Abnormal and Social Psychology*, 62, 461-464. PubMed
- Jones, Gerard *(2002). Killing monsters: why children need fantasy, super heroes and make-believe violence*. New York: Basic Books ISBN 978-0465036967
- Bureau M, Hirsch E, Vigevano F (2004). "Epilepsy and videogames". *Epilepsia* 45 Suppl 1: 24–6. PMID 14706041
- Web Content Accessibility Guidelines (WCAG) 2.0 - W3C Recommendation 11 December 2008: http://www.w3.org/TR/2008/REC-WCAG20-20081211/
- CERT: Email bombing and Spamming: http://www.cert.org/tech_tips/email_bombing_spamming.html
- Harold, Elliotte Rusty (27 May 2005). "Tip: Configure SAX parsers for secure processing". *IBM developerWorks*.

Research Articles

- So, R.H.Y. and Lo, W.T. (1999) "Cybersickness: An Experimental Study to Isolate the Effects of Rotational Scene Oscillations." Proceedings of IEEE Virtual Reality '99 Conference, March 13–17, 1999, Houston, Texas. Published by IEEE Computer Society, pp.237-241.
- G. Friedland, R. Sommer: Cybercasing the Joint: On the Privacy Implications of Geotagging, accepted for Usenix HotSec 2010 at the Usenix Security Conference, Washington DC, August 2010.

Exercises

1. Find examples in the media where multimedia data is attributed to a crime. Discuss the coverage with your fellow students.
2. Enter your name into a search engine. Count the number of occurences where a) the content is about you and b) the content is about you but not published by yourself. Find the oldest post about you.
3. List potential multimedia retrieval technologies (from this book and elsewhere) that could be used to invade privacy. Discuss possible counter measures.
4. Describe an inference chain over several websites that would compromise privacy. Discuss a second one that includes multimedia data.
5. Find your personal privacy sweet spot by discussing what would still be OK to be published about you and what would not.
6. Discuss the buffer overrun mentioned in the chapter in detail by taking one of the entropy compression algorithms from chapter XXX and creating a file that would expand into a very large file.